UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/684,488	10/04/2000	Bin Zhang	10992482-1	3131

7590   03/18/2003

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P.O. Box 272400
Fort Collins, CO   80528-9599

EXAMINER
HAMILTON, MONPLAISIR G

ART UNIT	PAPER NUMBER
2172	

DATE MAILED: 03/18/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 07-01)

Office Action Summary	Application No.	Applicant(s)
	09/684,488	ZHANG ET AL.
	Examiner	Art Unit
	Monplaisir G Hamilton	2172

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM
THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed
  after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any
  earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>24 December 2002</u> .

2a)☒ This action is **FINAL**.   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is
closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) <u>1-20</u> is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) <u>1-20</u> is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage
application from the International Bureau (PCT Rule 17.2(a)).
* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .

4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

1.   Claims 1-15 were pending. The communication filed on 12/24/02 added claims 16-20.

Claims 1-20 remain for examination.

Response to Arguments

2.   Applicant's arguments filed 12/24/2002 have been fully considered but they are not

persuasive.

Applicant argues, "the Guha reference is directed to a very different type of clustering (i.e., a

hierarchical clustering method). Guha appears to be directed to a class of hierarchical clustering

algorithms that operate in a very different manner as compared with the class of partitional clustering

algorithms of which the invention as claimed is directed.. Moreover, the c parameter of Guha is not the

same as the claimed size parameter. The c parameter represents a ''constant number of well scattered

points within the cluster," that "represents each cluster to be merged." (col. 6, lines 54 to 55) In sharp

contrast, the size parameter of the invention specifies the number of data points to be moved at one time

from one cluster to another cluster as claimed."

Examiner disagrees Guha explicitly discloses a partitioning scheme used for speeding up

his method (col 14, lines 50-55). This embodiment is essentially the same as the claimed

partitional clustering algorithm. Furthermore, in response to applicant's argument that the

references fails regarding the size parameter of the invention, examiner holds that Guha's c

parameter is essentially the same as the claimed size parameter.   The C parameter specifies the

number of data points that will be evaluated when deciding whether to merge a pair of clusters

(col 9, lines 30-68; col 10, lines 1-30). The merge procedure of Guha is essentially the same as the claimed move. In addition, although Guha does not explicitly set C to be the number of data points to be merged/moved, Guha does disclose situations when the size parameter (C) specifies the number of data points to be moved at one time from one cluster to another cluster e.g. when clustering initially begins the c is 1 (col 6, lines 55-65; col 7, lines 10-30), each data point in c is compared to every other cluster to find which cluster to merge/move the data points with/to (col 7, lines 35-40).

Applicant further argues, "the Guha reference fails to even identify or suggest the disadvantages or problems of prior art partitional clustering algorithms."

In response to applicant's argument that Guha is nonanalogous art, it has been held that a prior art reference must either be in the field of applicant's endeavor or, if not, then be reasonably pertinent to the particular problem with which the applicant was concerned, in order to be relied upon as a basis for rejection of the claimed invention. See *In re Oetiker*, 977 F.2d 1443, 24 USPQ2d 1443 (Fed. Cir. 1992). In this case, Guha is in the field of applicant's endeavor. Guha discloses this invention relates to the art of data mining, and in particular to a system for and method of clustering large databases (col 1, lines 5-10).

Applicant argues, "the Action states on page 3, "the teachings disclosed by Guha are essentially the same as the claimed limitations." Furthermore, the Action states, "the c parameter disclosed by Guha has the same functionality as the claimed size parameter, they both help in determining whether the datasets will be merged." It is respectfully submitted that these two statements are inaccurate. Applicant agrees with the Action that Guha fails to expressly disclose the steps of: (a) evaluating subsets of data

points in each cluster for moving into every other cluster by using a predetermined metric; wherein the number of data points in the subset is specified by the size parameter..... the Action admits that the specific claim limitations added by each of the dependent claims are not expressly disclosed by the Guha reference. The Action then continues "However, the teachings disclosed by Guha are essentially the same as the claimed invention. It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases. (Col. 1, lines 5-10).

Applicant's arguments fail to comply with 37 CFR 1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references.

Claim Objections

3.   Claim 16 is objected to because of the following informalities: "based on based on" in limitation should be "based on". Appropriate correction is required.

Claim 17 is objected to because of the following informalities: m_u, m_j, m_i n_i and n_j are not defined. Appropriate correction is required.

Claims 19-20 are objected to because of the following informalities: u should be U. Appropriate correction is required

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

4.   Claims 1-10 and 16-20 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Regarding Claims 1-10 and 16-20:

Claims 1-10 and 16-20 are related to method of clustering data wherein data belonging a cluster is compared against data points in all other clusters to decide whether the data should be moved to a different cluster. The claimed invention is not within the technological arts because these claims are related to a process that consists solely of the manipulation of an abstract idea. The method as claimed can be performed by a person using a paper and pen. Furthermore, the applicant's disclosure does not relate the claimed method to functional descriptive material (computer-readable medium).

Regarding Claims 11-15:

Claims 11-15 describe a system for implementing the methods of Claims 1-10 and 16-20. The claimed invention is not within the technological arts because these claims are related to software per se, a system without hardware.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in-
(1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effect under this subsection of a national application published under section 122(b) only if the international application designating the United States was published under Article 21(2)(a) of such treaty in the English language; or
(2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that a patent shall not be deemed filed in the United States for the purposes of this subsection based on the filing of an international application filed under the treaty defined in section 351(a).

5.   Claim 16 is rejected under 35 U.S.C. 102(e) as being anticipated by US 6269376 issued to Dhillon et al.

A partitional method for clustering comprising the steps of:

a) receiving K initial center positions; wherein each center is denoted by m_k wherein k=1,...K; and wherein K is the number of clusters (col 5, lines 1-20)

b) receiving a plurality of data points and partitioning the data points into a plurality of clusters based on a relationship between the data point and the center point of a respective cluster (col 1, lines 60-65);

c) at least two data points in a first partition Si are simultaneously evaluated for moving to every other partition (col 6, lines 30-45; col 7, lines 38-43); wherein a subsets of U points evaluated for moving (col 7, lines 38-43); wherein and index j is utilized to represent the partition to which a data point x currently belongs (col 6, lines 40-45), and the index i is utilized to represent the partition that is currently being evaluated for a potential move to which the data point x can be moved (col 6, lines 45-50; col 7, lines 1-5).

Claim Rejections - 35 USC § 103

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

6.   Claims 1-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over US 6092072 issued to Guha.

Referring to Claims 1-15:

Please see prior office action dated 9/11/2002.

7.   Claims 17-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over US 6269376 issued to Dhillon et al further in view of US 6092072 issued to Guha.

Referring to Claims 17:

Dhillon discloses the limitations as discussed in Claim 16 above.

Dhillon discloses the use of a metric to calculate the distance between a point and a center.

Dhillon does not explicitly discloses the claimed "predetermined metric is employed for evaluating whether a set of data points U should be moved from the current partition to a second partition; wherein the predetermined metric includes expression:

$$\frac{n_i}{n_i - |U|}|m_U - m_i|^2 - \frac{n_j}{n_j + |U|}|m_U - m_j|^2$$

When the predetermined metric is in a first predetermined relationship with a

predetermined value, moving the set of data points U is moved from a current partition SI to a

second partition Sj"

Guha discloses the distance measure or square error criterion is used on cluster centers to

decide whether to merge the clusters (col 1, lines 29-65). Furthermore when $U=n_i$ the claimed

expression returns the square error criterion. Therefore at the time the invention was made, it

would have been obvious to a person of ordinary skill in the art to modify Dhillon to have a

method of merging/moving cluster based on a square error criterion. One of ordinary skill in the

art would have been motivated to do this because it would provide a method for clustering large

data sets (col 1, lines 15-20).

Referring to Claim 18:

Dhillon in view of Guha discloses the limitations as discussed in Claim 17 above. Guha

further discloses merging/moving the set of data points from a current position to a second

partition includes the steps of: updating the count of the current partition and the count of the

second partition; and updating the center of the current position and the center of the second

partition (col 7, lines 22-26; and col 8, lines 25-35).

Referring to Claim 19:

Dhillon in view of Guha discloses the limitations as discussed in Claim 18 above. Guha

further discloses updating the counts of the two partitions utilizes the following expressions:

$n_i=n_i-|u|$, and $n_j=n_j+|u|$; wherein U is the subset of data points (U is a subset of Si) being

evaluated for the move, |U| is the size of U that is specified by the size parameter, m_U is the geometric center of U, m_i and m_j are the centers of the clusters and n_i and n_j are the counts of the clusters (col 7, lines 22-26; and col 8, lines 25-35).

Referring to Claim 20:

Dhillon in view of Guha discloses the limitations as discussed in Claim 18 above. Guha further discloses updating the centers of these two partitions utilizes the following expressions:

Mi=(nI*mi-mu)/(ni-u), and mj=(nj*mj+mu)/(ni+u)

wherein U is the subset of data points (U is a subset of Si) being evaluated for the move, U is the size of U that is specified by the size parameter , mU is the geometric center of U, mi and mj are the centers of the clusters and ni and nj are the counts of the clusters (col 7, lines 22-26; and col 8, lines 25-35).

Prior Art

8.   The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

US 6421668 issued to Yakhini, Zohar H et al. Yakhini discloses a method and system for applying arbitrary similarity metrics to data entities in order to partition the entities into subsets of related entities. The method and system iteratively construct successive subsets, during construction of each subset adding candidate entities, not yet assigned to a subset, with high affinities toward the subset and removing entities previously assigned to the subset for which the affinities toward the subset have decreased.

US 5619709 issued to Caid, William R. et al. Caid discloses context vectors represent

conceptual relationships among information items by quantitative means. A neural network

operates on a training corpus of records to develop relationship-based context vectors based on

word proximity and co-importance using a technique of "windowed co-occurrence".

Relationships among context vectors are deterministic, so that a context vector set has one

logical solution, although it may have a plurality of physical solutions. No human knowledge,

thesaurus, synonym list, knowledge base, or conceptual hierarchy, is required. Summary

vectors of records may be clustered to reduce searching time, by forming a tree of clustered

nodes. Once the context vectors are determined, records may be retrieved using a query

interface that allows a user to specify content terms, Boolean terms, and/or document feedback.

Final Rejection

9.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within TWO

MONTHS of the mailing date of this final action and the advisory action is not mailed until after

the end of the THREE-MONTH shortened statutory period, then the shortened statutory period

will expire on the date the advisory action is mailed, and any extension fee pursuant to 37

CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event,

however, will the statutory period for reply expire later than SIX MONTHS from the mailing
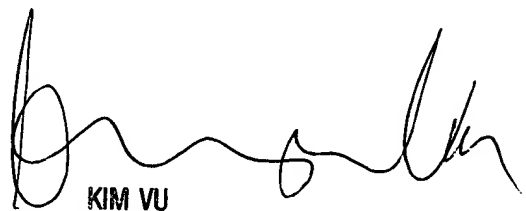
date of this final action.

Conclusion

10.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Monplaisir G Hamilton whose telephone number is 1703-305-

5116. The examiner can normally be reached on Monday - Friday (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Kim Y Vu can be reached on 1703-305-4393. The fax phone numbers for the

organization where this application or proceeding is assigned are 1703-746-7239 for regular

communications and 1703-746-7238 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding

should be directed to the receptionist whose telephone number is 1703-305-3900.

Monplaisir Hamilton
February 27, 2003

KIM VU
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100